Universität Stuttgart

Institut für
Maschinelle
Sprachverarbeitung

Janis Pagel
Nils Reiter
Ina Rösiger
Sarah Schulz

A Unified Text Annotation Workflow for Diverse Goals

Why do we annotate?

- Empirical validation of theories

- Data creation

Why do we annotate?

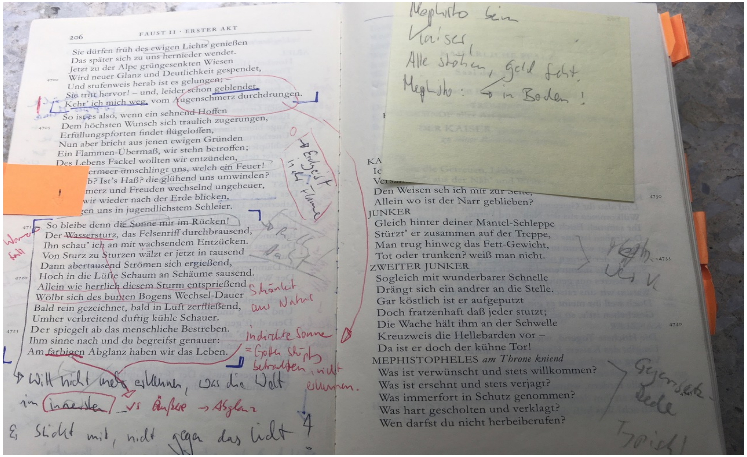
- Empirical validation of theories
  - Discovering phenomena not covered by a theory
  - Strengthening definitions in a theory
    - Often confused categories might be overlapping or at least unclear
  - Uncovering implicit assumptions
- Data creation

Why do we annotate?

- Empirical validation of theories
  - Discovering phenomena not covered by a theory
  - Strengthening definitions in a theory
    - Often confused categories might be overlapping or at least unclear
  - Uncovering implicit assumptions
- Data creation
  - Manually annotated data can be analysed
    - Which categories are how frequent in what context?
  - Automatic tools can be evaluated
    - How well do machines do this task?
  - Supervised tools can be trained

Analog annotation …

- Ideas attached to spans of text
  - Sometimes fuzzy text spans



206 FAUST II · ERSTER AKT

Sie dürfen früh des ewigen Lichts genießen
Das später sich zu uns hernieder wendet.
Jetzt zu der Alpe grüngesenkten Wiesen
Wird neuer Glanz und Deutlichkeit gespendet,
Und stufenweis herab ist es gelungen; –
Sie tritt hervor! – und, leider schon geblendet,
Kehr' ich mich weg, vom Augenschmerz durchdrungen.

So ist es also, wenn ein sehnend Hoffen
Dem höchsten Wunsch sich traulich zugerungen,
Erfüllungspforten findet flügeloffen,
Nun aber bricht aus jenen ewigen Gründen
Ein Flammen-Übermaß, wir stehn betroffen;
Des Lebens Fackel wollten wir entzünden,
…ermeer umschlingt uns, welch ein Feuer!
…b? Ist's Haß? die glühend uns umwinden?
…merz und Freuden wechselnd ungeheuer,
…wir wieder nach der Erde blicken,
…gen uns in jugendlichstem Schleier.

So bleibe denn die Sonne mir im Rücken!
Der Wassersturz, das Felsenriff durchbrausend,
Ihn schau' ich an mit wachsendem Entzücken.
Von Sturz zu Sturzen wälzt er jetzt in tausend
Dann abertausend Strömen sich ergießend,
Hoch in die Lüfte Schaum an Schäume sausend.
Allein wie herrlich diesem Sturm entsprießend
Wölbt sich des bunten Bogens Wechsel-Dauer
Bald rein gezeichnet, bald in Luft zerfließend,
Umher verbreitend duftig kühle Schauer.
Der spiegelt ab das menschliche Bestreben.
Ihm sinne nach und du begreifst genauer:
Am farbigen Abglanz haben wir das Leben.

!

Erdgeist in der Flamme

Wasserfall

Schönheit aus Natur

Indirekte Sonne = Gottes Schöpfung betrachten, nicht erkennen.

Will nicht mehr erkennen, was die Welt im Innersten – vs Äußere → Abglanz

Er steht mit, nicht gegen das Licht!

Mephisto beim Kaiser,
Alle stöhnen, Geld fehlt.
Mephisto: → in Boden!

KA…
Ic…
Versa…
Den Weisen seh ich mir zur Seite,
Allein wo ist der Narr geblieben?
JUNKER
Gleich hinter deiner Mantel-Schleppe
Stürzt' er zusammen auf der Treppe,
Man trug hinweg das Fett-Gewicht,
Tot oder trunken? weiß man nicht.
ZWEITER JUNKER
Sogleich mit wunderbarer Schnelle
Drängt sich ein andrer an die Stelle.
Gar köstlich ist er aufgeputzt
Doch fratzenhaft daß jeder stutzt;
Die Wache hält ihm an der Schwelle
Kreuzweis die Hellebarden vor –
Da ist er doch der kühne Tor!
MEPHISTOPHELES *am Throne kniend*
Was ist verwünscht und stets willkommen?
Was ist ersehnt und stets verjagt?
Was immerfort in Schutz genommen?
Was hart gescholten und verklagt?
Wen darfst du nicht herbeiberufen?

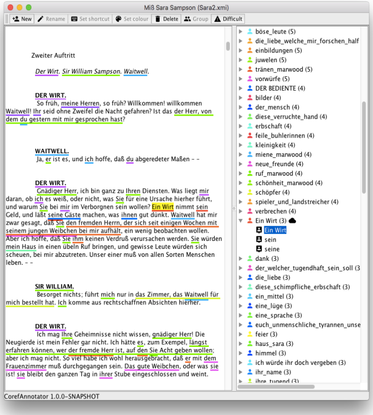
Meph. [illegible]

Gegensatz Rede

Typisch!

…and digital annotation

- Explicit assignment of categories to text spans
  - Text spans are explicitly bounded (begin, end)



Zweiter Auftritt

Der Wirt. Sir William Sampson. Waitwell.

DER WIRT.
So früh, meine Herren, so früh? Willkommen! willkommen Waitwell! Ihr seid ohne Zweifel die Nacht gefahren? Ist das der Herr, von dem du gestern mit mir gesprochen hast?

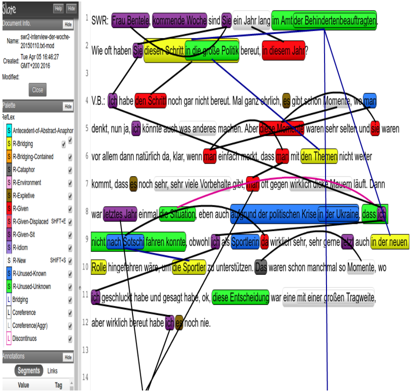
WAITWELL.
Ja, er ist es, und ich hoffe, daß du abgeredeter Maßen – –

DER WIRT.
Gnädiger Herr, ich bin ganz zu Ihren Diensten. Was liegt mir daran, ob ich es weiß, oder nicht, was Sie für eine Ursache hierher führt, und warum Sie bei mir im Verborgnen sein wollen? Ein Wirt nimmt sein Geld, und läßt seine Gäste machen, was ihnen gut dünkt. Waitwell hat mir zwar gesagt, daß Sie den fremden Herrn, der sich seit einigen Wochen mit seinem jungen Weibchen bei mir aufhält, ein wenig beobachten wollen. Aber ich hoffe, daß Sie ihm keinen Verdruß verursachen werden. Sie würden mein Haus in einen übeln Ruf bringen, und gewisse Leute würden sich scheuen, bei mir abzutreten. Unser einer muß von allen Sorten Menschen leben. – –

SIR WILLIAM.
Besorget nichts; führt mich nur in das Zimmer, das Waitwell für mich bestellt hat. Ich komme aus rechtschaffnen Absichten hierher.

DER WIRT.
Ich mag Ihre Geheimnisse nicht wissen, gnädiger Herr! Die Neugierde ist mein Fehler gar nicht. Ich hätte es, zum Exempel, längst erfahren können, wer der fremde Herr ist, auf den Sie Acht geben wollen; aber ich mag nicht. So viel habe ich wohl herausgebracht, daß er mit dem Frauenzimmer muß durchgegangen sein. Das gute Weibchen, oder was sie ist! sie bleibt den ganzen Tag in ihrer Stube eingeschlossen und weint.

- böse_leute (5)
- die_liebe_welche_mir_forschen_half
- einbildungen (5)
- juwelen (5)
- tränen_marwood (5)
- vorwürfe (5)
- DER BEDIENTE (4)
- bilder (4)
- der_mensch (4)
- diese_verruchte_hand (4)
- erbschaft (4)
- feile_buhlerinnen (4)
- kleinigkeit (4)
- miene_marwood (4)
- neue_freunde (4)
- ruf_marwood (4)
- schönheit_marwood (4)
- schöpfer (4)
- spieler_und_landstreicher (4)
- verbrechen (4)
- Ein Wirt (3)
  - Ein Wirt
  - sein
  - seine
- dank (3)
- der_welcher_tugendhaft_sein_soll (3
- die_liebe (3)
- diese_schimpfliche_erbschaft (3)
- ein_mittel (3)
- eine_lüge (3)
- eine_sprache (3)
- euch_unmenschliche_tyrannen_unse
- feier (3)
- haus_sara (3)
- himmel (3)
- ich würde ihr doch vergeben (3)
- ihr_name (3)
- ihre_tugend (3)



Name: swr2-interview-der-woche-20150110.txt-mod
Created: Tue Apr 05 16:46:27 GMT+200 2016
Modified:

SWR: Frau Bentele, kommende Woche sind Sie ein Jahr lang im Amt der Behindertenbeauftragten.

Wie oft haben Sie diesen Schritt in die große Politik bereut, in diesem Jahr?

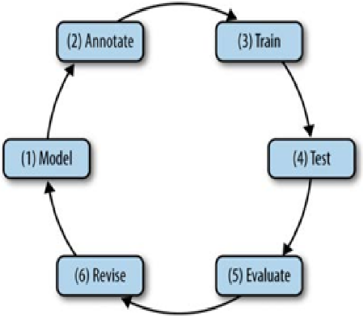
V.B.: Ich habe den Schritt noch gar nicht bereut. Mal ganz ehrlich, es gibt schon Momente, wo man denkt, nun ja, ich könnte auch was anderes machen. Aber diese Momente waren sehr selten und sie waren vor allem dann natürlich da, klar, wenn man einfach merkt, dass man mit den Themen nicht weiter kommt, dass es noch sehr, sehr viele Vorbehalte gibt, man oft gegen wirklich dicke Mauern läuft. Dann war letztes Jahr einmal die Situation, eben auch aufgrund der politischen Krise in der Ukraine, dass ich nicht nach Sotschi fahren konnte, obwohl ich als Sportlerin da wirklich sehr, sehr gerne jetzt auch in der neuen Rolle hingefahren wäre, um die Sportler zu unterstützen. Das waren schon manchmal so Momente, wo ich geschluckt habe und gesagt habe, ok, diese Entscheidung war eine mit einer großen Tragweite, aber wirklich bereut habe ich es noch nie.

Circles

- **Annotation (Circle)**
  - Well known in Computational Linguistics
  - Automation-centered
  - “Enrich with knowledge”

Circles

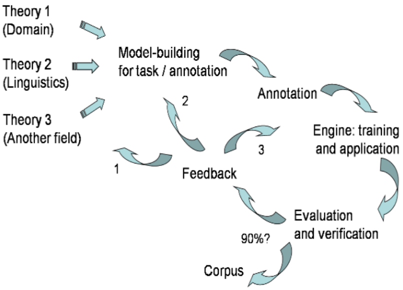
- **Annotation (Circle)**
  - Well known in Computational Linguistics
  - Automation-centered
  - “Enrich with knowledge”



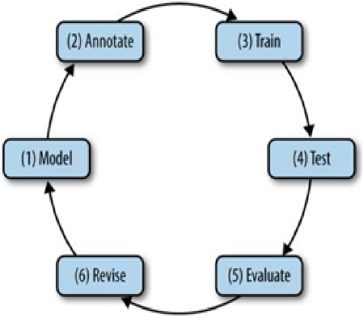
Pustejovsky and Stubbs (2012)

Circles

- **Annotation (Circle)**
  - Well known in Computational Linguistics
  - Automation-centered
  - “Enrich with knowledge”



Hovy and Lavid (2010)



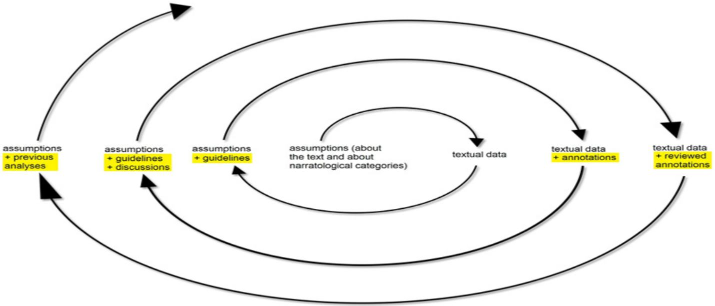
Pustejovsky and Stubbs (2012)

Circles

- **Hermeneutic Circle**
  - Well known in Humanities
  - Interpretation-centered
  - “Retrieve knowledge”

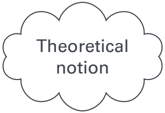
Circles

- **Hermeneutic Circle**
  - Well known in Humanities
  - Interpretation-centered
  - “Retrieve knowledge”


assumptions + previous analyses
assumptions + guidelines + discussions
assumptions + guidelines
assumptions (about the text and about narratological categories)
textual data
textual data + annotations
textual data + reviewed annotations

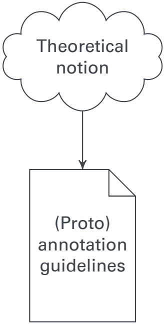
Gius and Jacke (2017)

Workflow

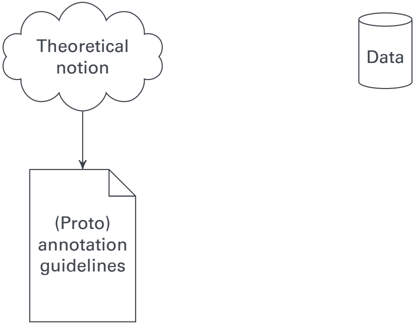


Theoretical notion

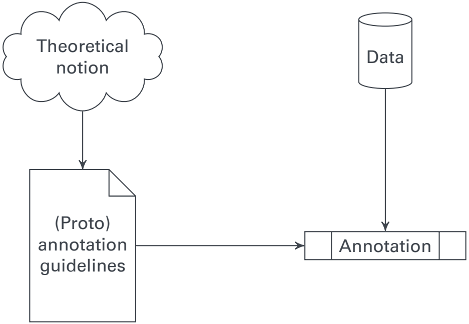
Workflow



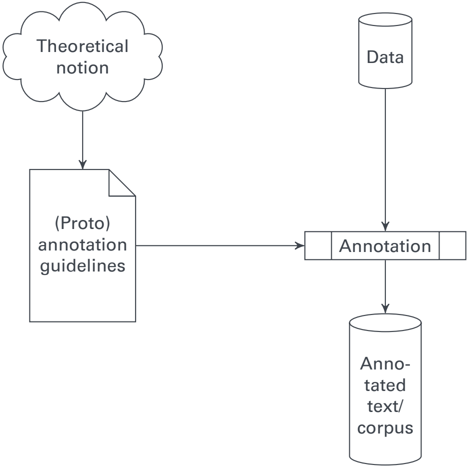
Workflow



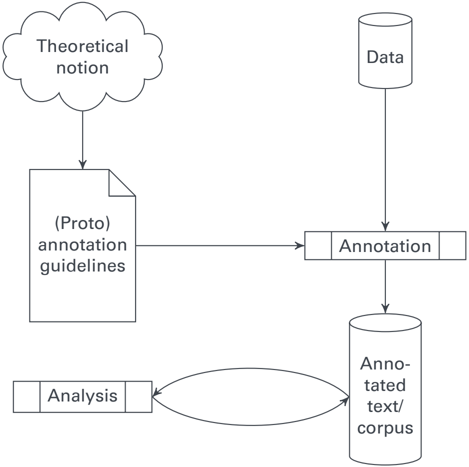
Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation

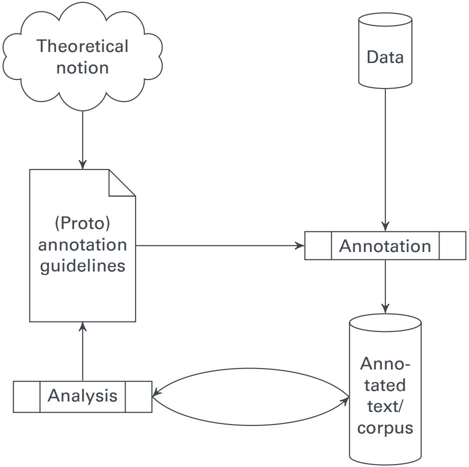
Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno-tated text/ corpus

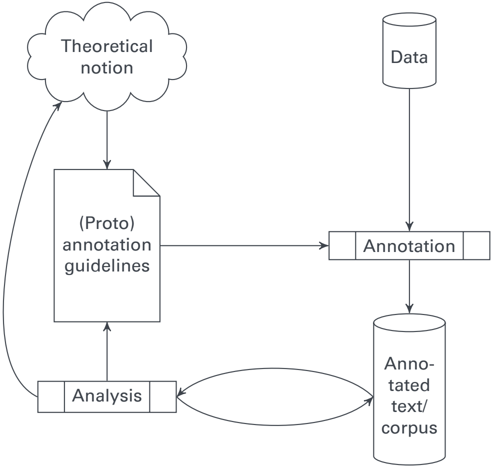
Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno-tated text/ corpus
Analysis

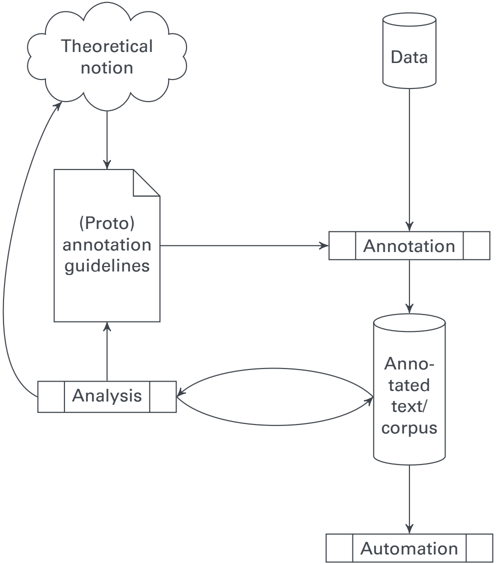
Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno- tated text/ corpus
Analysis

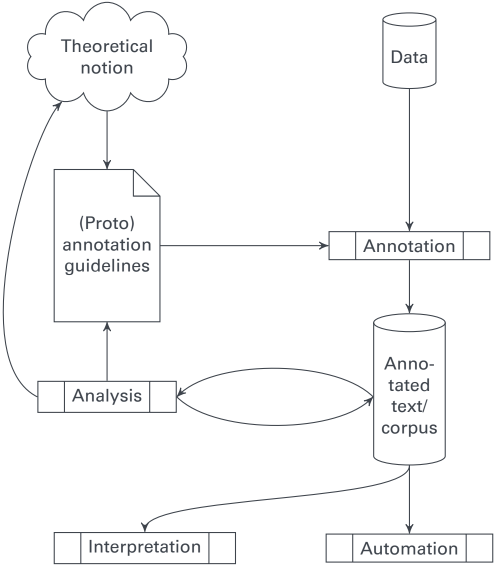
Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno-tated text/ corpus
Analysis

Workflow


Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno-tated text/ corpus
Analysis
Automation

Workflow

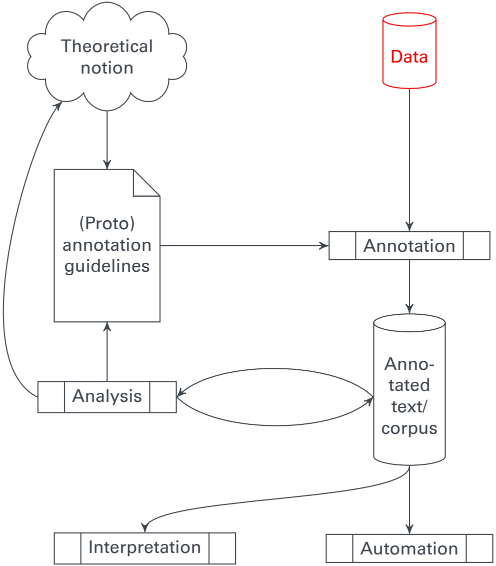

Theoretical notion
Data
(Proto) annotation guidelines
Annotation
Anno-tated text/ corpus
Analysis
Interpretation
Automation

Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions

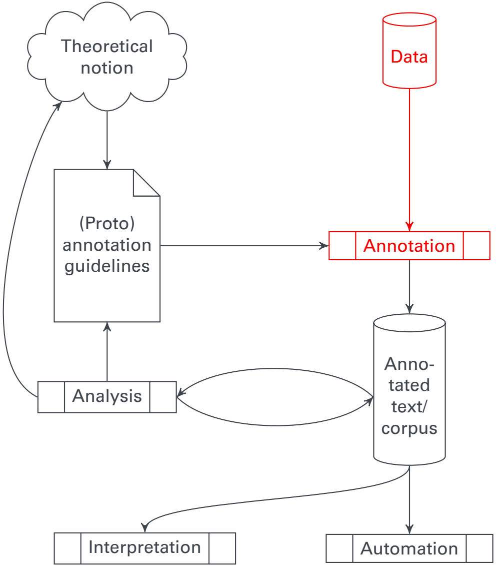
Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions



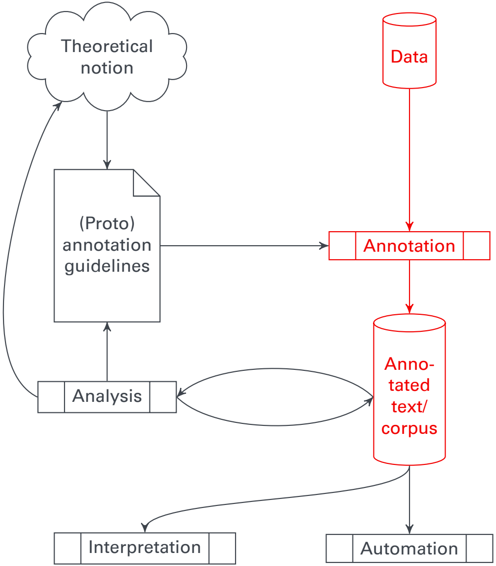
Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions



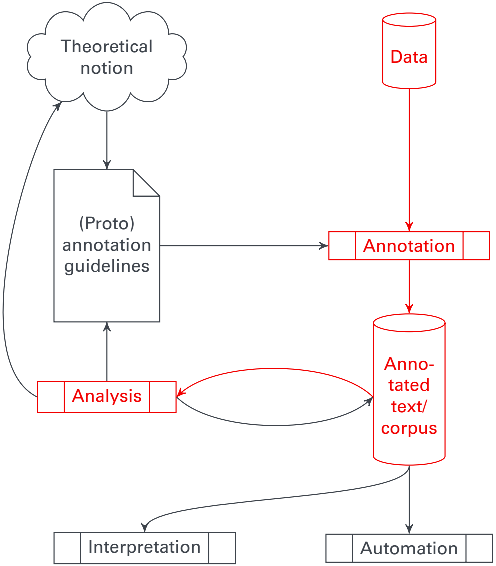
Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions



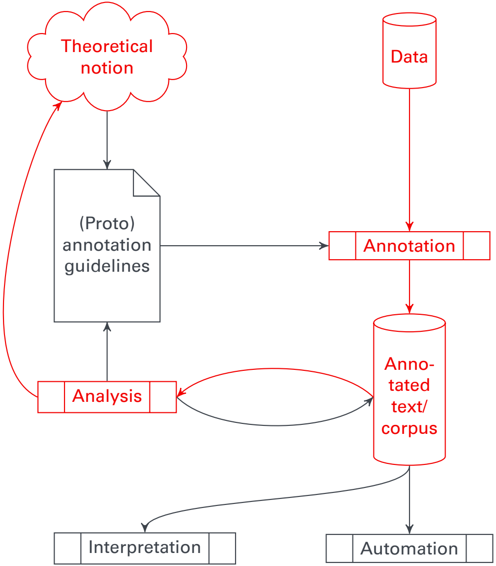
Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions



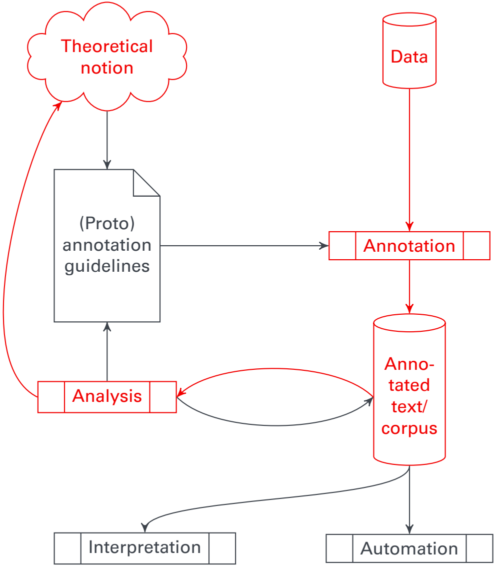
Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions

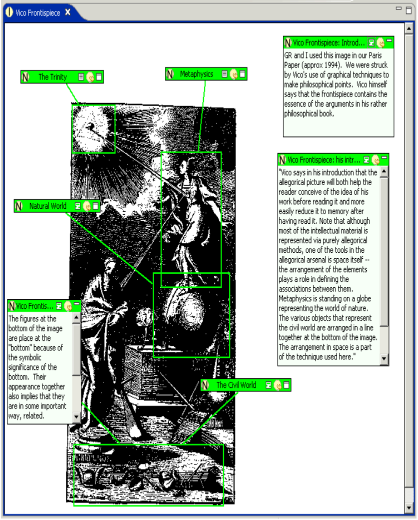


Goals of Annotation

- **Exploratory**
  - Mostly note-taking
  - Semi-organized
  - Humanities centered
  - Ideally completely free of presuppositions
- Examples
  - *Pliny* (Bradley, 2008)
  - *3DH* (Kleymann, Meister, and Stange, 2018)



Pliny



Vico Frontispiece: Introd…

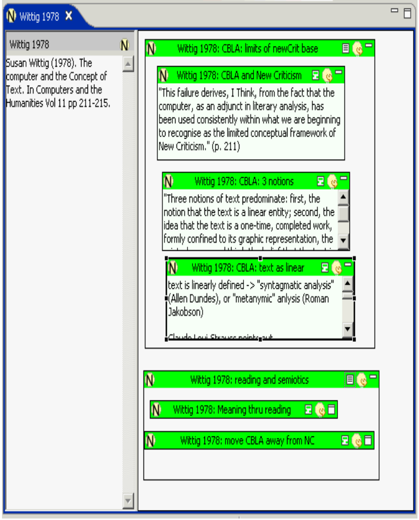
GR and I used this image in our Paris Paper (approx 1994). We were struck by Vico's use of graphical techniques to make philosophical points. Vico himself says that the frontispiece contains the essence of the arguments in his rather philosophical book.

Vico Frontispiece: his intr…

"Vico says in his introduction that the allegorical picture will both help the reader conceive of the idea of his work before reading it and more easily reduce it to memory after having read it. Note that although most of the intellectual material is represented via purely allegorical methods, one of the tools in the allegorical arsenal is space itself -- the arrangement of the elements plays a role in defining the associations between them. Metaphysics is standing on a globe representing the world of nature. The various objects that represent the civil world are arranged in a line together at the bottom of the image. The arrangement in space is a part of the technique used here."

Vico Frontis…

The figures at the bottom of the image are place at the "bottom" because of the symbolic significance of the bottom. Their appearance together also implies that they are in some important way, related.



Wittig 1978

Susan Wittig (1978). The computer and the Concept of Text. In Computers and the Humanities Vol 11 pp 211-215.

Wittig 1978: CBLA: limits of newCrit base

Wittig 1978: CBLA and New Criticism

"This failure derives, I Think, from the fact that the computer, as an adjunct in literary analysis, has been used consistently within what we are beginning to recognise as the limited conceptual framework of New Criticism." (p. 211)

Wittig 1978: CBLA: 3 notions

"Three notions of text predominate: first, the notion that the text is a linear entity; second, the idea that the text is a one-time, completed work, formly confined to its graphic representation, the

Wittig 1978: CBLA: text as linear

text is linearly defined -> "syntagmatic analysis" (Allen Dundes), or "metanymic" anlysis (Roman Jakobson)

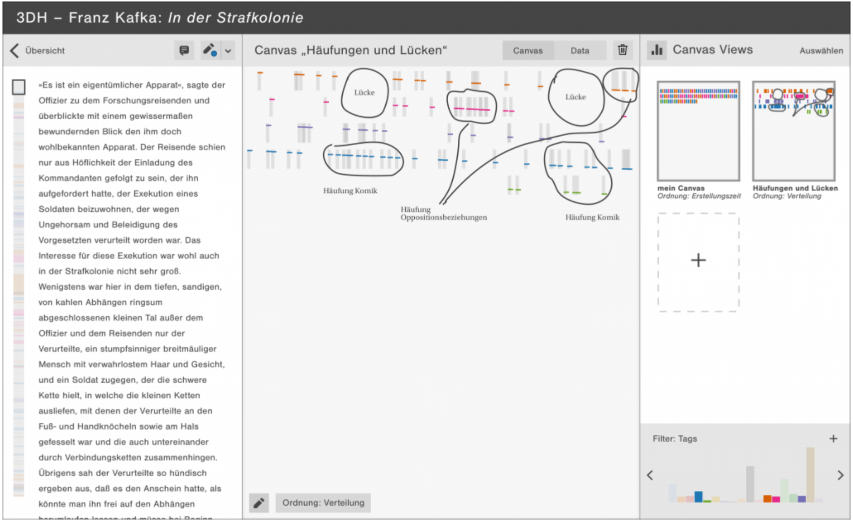
Claude Loui Strauss points out

Wittig 1978: reading and semiotics

Wittig 1978: Meaning thru reading

Wittig 1978: move CBLA away from NC

3DH



3DH – Franz Kafka: *In der Strafkolonie*

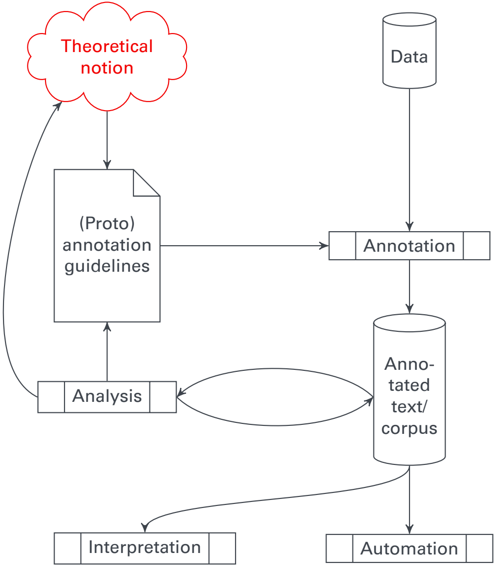
»Es ist ein eigentümlicher Apparat«, sagte der Offizier zu dem Forschungsreisenden und überblickte mit einem gewissermaßen bewundernden Blick den ihm doch wohlbekannten Apparat. Der Reisende schien nur aus Höflichkeit der Einladung des Kommandanten gefolgt zu sein, der ihn aufgefordert hatte, der Exekution eines Soldaten beizuwohnen, der wegen Ungehorsam und Beleidigung des Vorgesetzten verurteilt worden war. Das Interesse für diese Exekution war wohl auch in der Strafkolonie nicht sehr groß. Wenigstens war hier in dem tiefen, sandigen, von kahlen Abhängen ringsum abgeschlossenen kleinen Tal außer dem Offizier und dem Reisenden nur der Verurteilte, ein stumpfsinniger breitmäuliger Mensch mit verwahrlostem Haar und Gesicht, und ein Soldat zugegen, der die schwere Kette hielt, in welche die kleinen Ketten ausliefen, mit denen der Verurteilte an den Fuß- und Handknöcheln sowie am Hals gefesselt war und die auch untereinander durch Verbindungsketten zusammenhingen. Übrigens sah der Verurteilte so hündisch ergeben aus, daß es den Anschein hatte, als könnte man ihn frei auf den Abhängen herumlaufen lassen und müsse bei Beginn

Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts

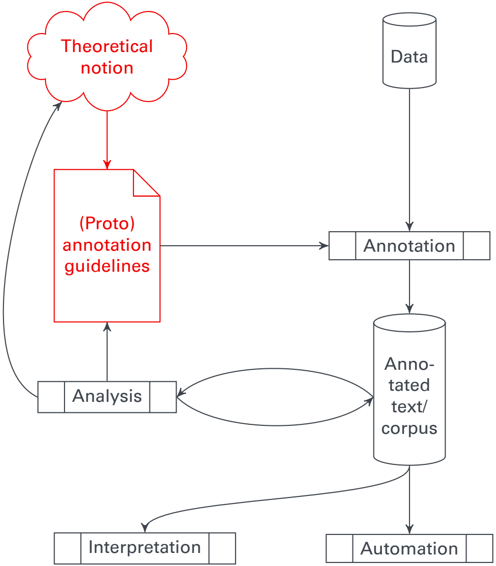
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



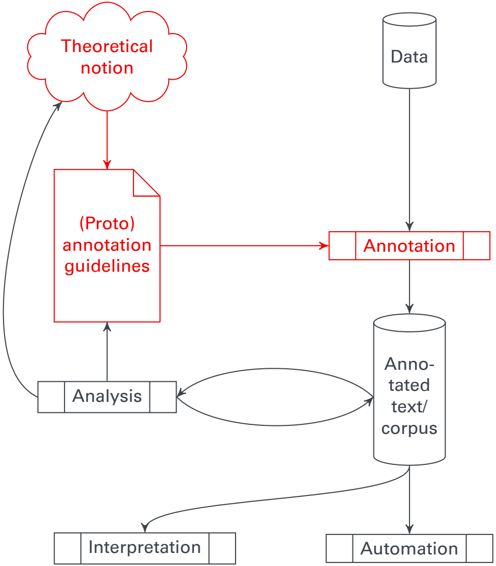
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



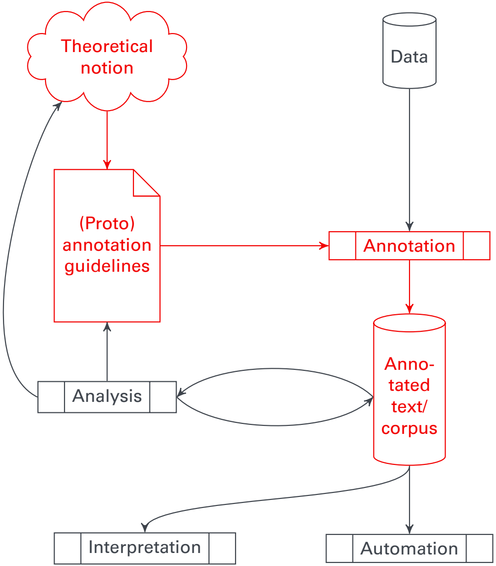
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



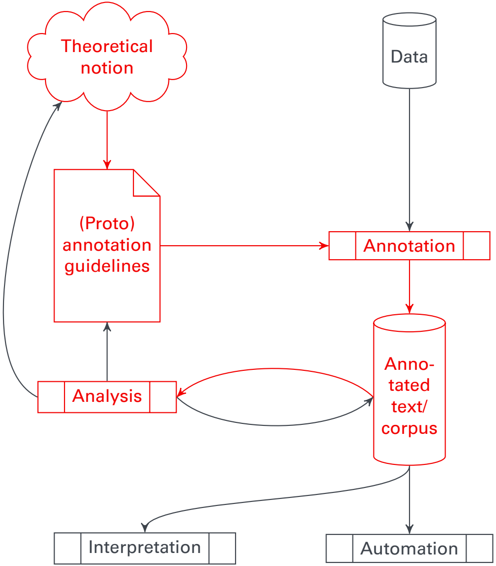
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



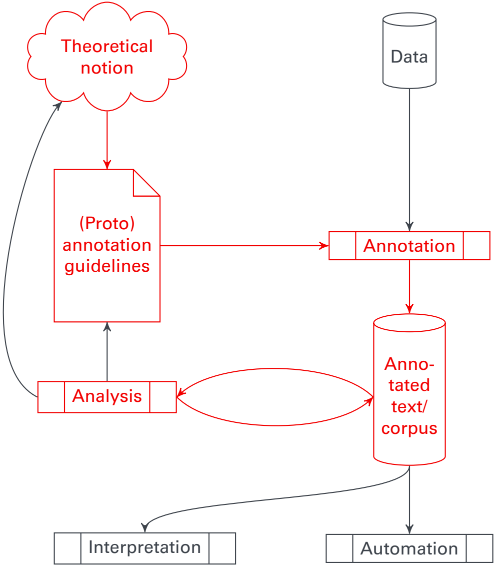
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



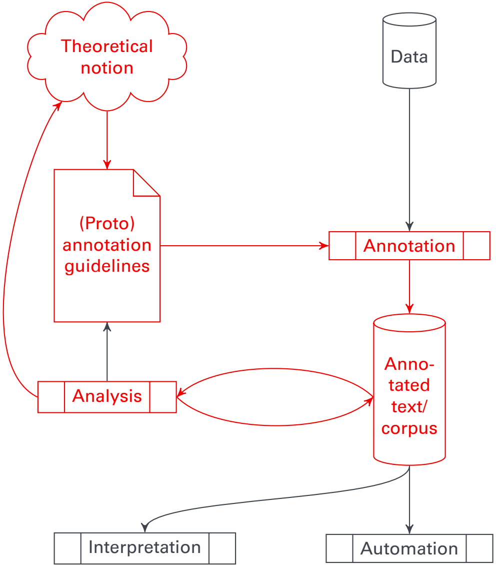
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts



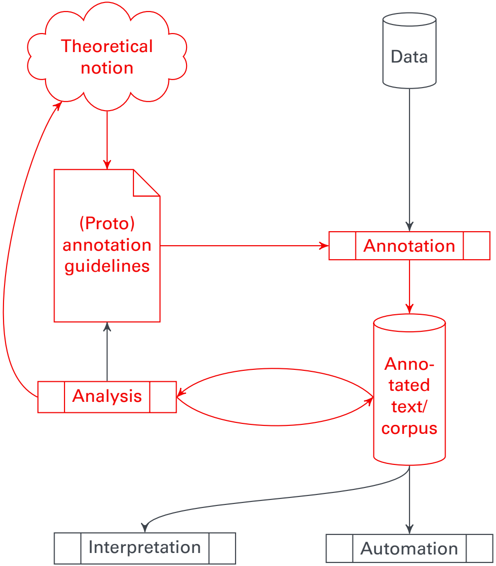
Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts

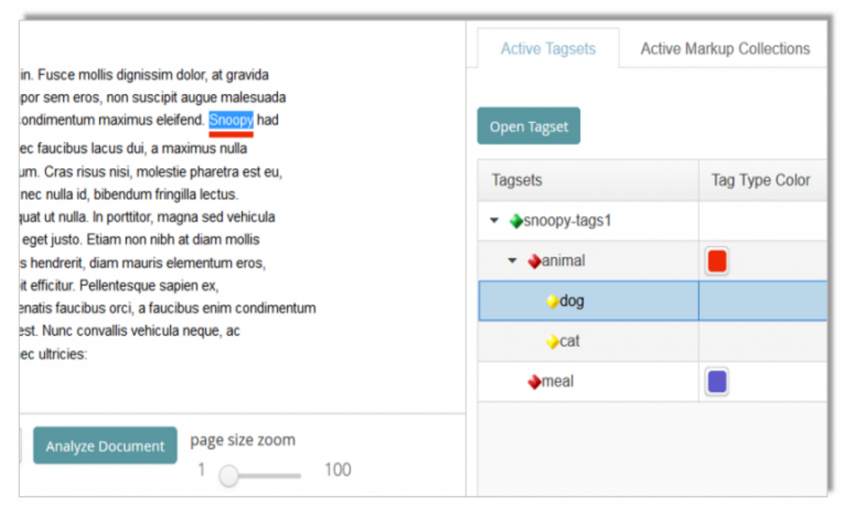


Goals of Annotation

- **Conceptualizing**
  - Sharpen theoretic notions
  - Find disagreements of theoretical claims on concrete texts
- Examples
  - *heureCLÉA* (Bögel et al., 2015)
  - *QuaDramA* (Rösiger, Schulz, and Reiter, 2018)



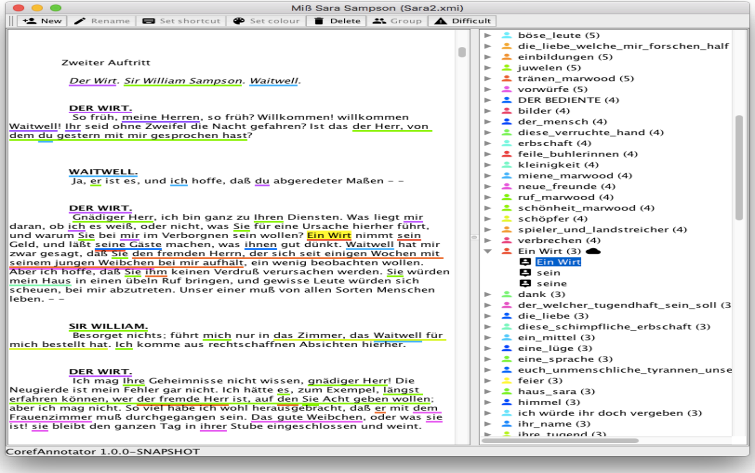
heureCLÉA (Catma)



in. Fusce mollis dignissim dolor, at gravida
por sem eros, non suscipit augue malesuada
ondimentum maximus eleifend. Snoopy had
ec faucibus lacus dui, a maximus nulla
um. Cras risus nisi, molestie pharetra est eu,
nec nulla id, bibendum fringilla lectus.
uat ut nulla. In porttitor, magna sed vehicula
eget justo. Etiam non nibh at diam mollis
s hendrerit, diam mauris elementum eros,
it efficitur. Pellentesque sapien ex,
enatis faucibus orci, a faucibus enim condimentum
est. Nunc convallis vehicula neque, ac
ec ultricies:

Tagsets	Tag Type Color
▾ snoopy-tags1	
▾ animal	🟥
dog	
cat	
meal	🟪

QuaDramA



Zweiter Auftritt

Der Wirt. Sir William Sampson. Waitwell.

DER WIRT.
So früh, meine Herren, so früh? Willkommen! willkommen Waitwell! Ihr seid ohne Zweifel die Nacht gefahren? Ist das der Herr, von dem du gestern mit mir gesprochen hast?

WAITWELL.
Ja, er ist es, und ich hoffe, daß du abgeredeter Maßen – –

DER WIRT.
Gnädiger Herr, ich bin ganz zu Ihren Diensten. Was liegt mir daran, ob ich es weiß, oder nicht, was Sie für eine Ursache hierher führt, und warum Sie bei mir im Verborgnen sein wollen? Ein Wirt nimmt sein Geld, und läßt seine Gäste machen, was ihnen gut dünkt. Waitwell hat mir zwar gesagt, daß Sie den fremden Herrn, der sich seit einigen Wochen mit seinem jungen Weibchen bei mir aufhält, ein wenig beobachten wollen. Aber ich hoffe, daß Sie ihm keinen Verdruß verursachen werden. Sie würden mein Haus in einen übeln Ruf bringen, und gewisse Leute würden sich scheuen, bei mir abzutreten. Unser einer muß von allen Sorten Menschen leben. – –

SIR WILLIAM.
Besorget nichts; führt mich nur in das Zimmer, das Waitwell für mich bestellt hat. Ich komme aus rechtschaffnen Absichten hierher.

DER WIRT.
Ich mag Ihre Geheimnisse nicht wissen, gnädiger Herr! Die Neugierde ist mein Fehler gar nicht. Ich hätte es, zum Exempel, längst erfahren können, wer der fremde Herr ist, auf den Sie Acht geben wollen; aber ich mag nicht. So viel habe ich wohl herausgebracht, daß er mit dem Frauenzimmer muß durchgegangen sein. Das gute Weibchen, oder was sie ist! sie bleibt den ganzen Tag in ihrer Stube eingeschlossen und weint.

- böse_leute (5)
- die_liebe_welche_mir_forschen_half
- einbildungen (5)
- juwelen (5)
- tränen_marwood (5)
- vorwürfe (5)
- DER BEDIENTE (4)
- bilder (4)
- der_mensch (4)
- diese_verruchte_hand (4)
- erbschaft (4)
- feile_buhlerinnen (4)
- kleinigkeit (4)
- miene_marwood (4)
- neue_freunde (4)
- ruf_marwood (4)
- schönheit_marwood (4)
- schöpfer (4)
- spieler_und_landstreicher (4)
- verbrechen (4)
- Ein Wirt (3)
  - Ein Wirt
  - sein
  - seine
- dank (3)
- der_welcher_tugendhaft_sein_soll (3
- die_liebe (3)
- diese_schimpfliche_erbschaft (3)
- ein_mittel (3)
- eine_lüge (3)
- eine_sprache (3)
- euch_unmenschliche_tyrannen_unse
- feier (3)
- haus_sara (3)
- himmel (3)
- ich würde ihr doch vergeben (3)
- ihr_name (3)
- ihre_tugend (3)

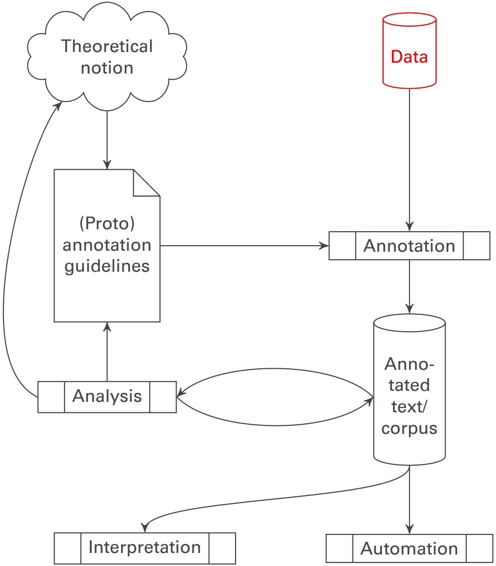
CorefAnnotator 1.0.0-SNAPSHOT

Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation

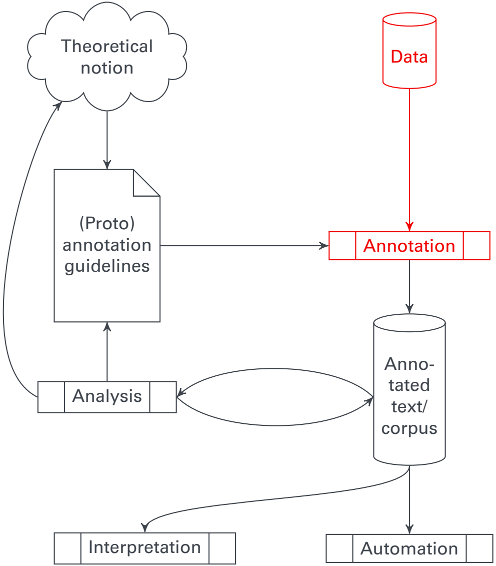
Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation



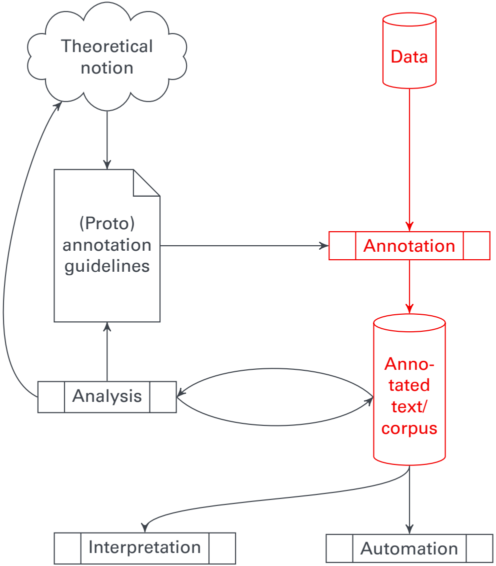
Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation



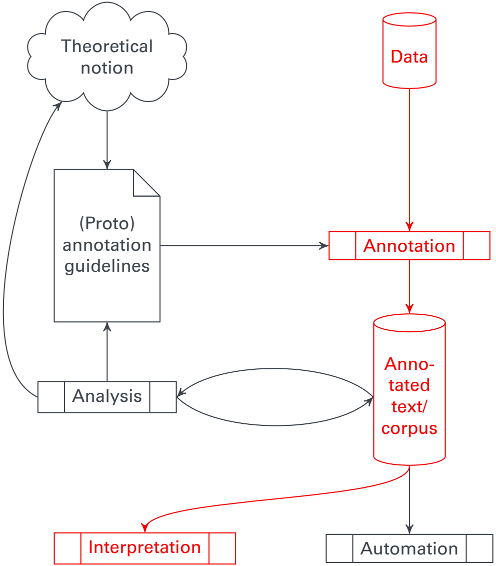
Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation



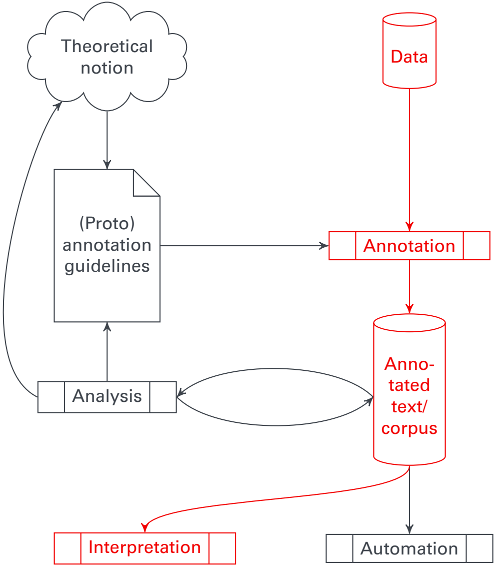
Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation

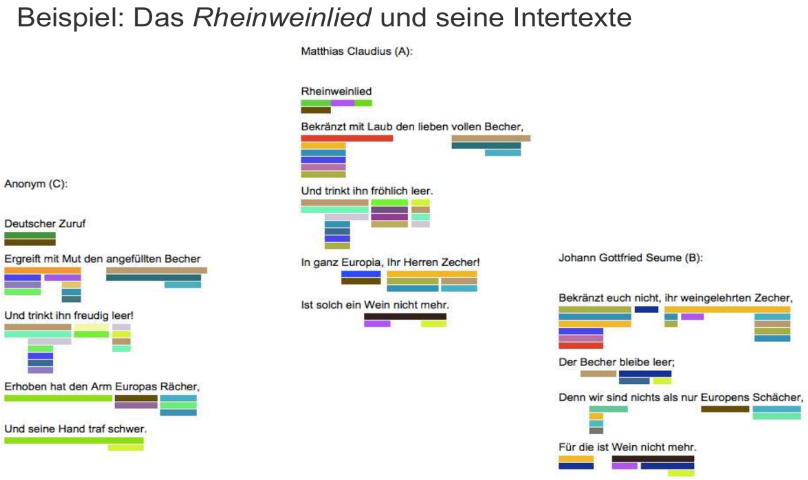


Goals of Annotation

- **Explicating**
  - Structure text in an observable way
  - Helps interpretation
- Examples
  - Nantke and Schlupkothen (2018)



Nantke and Schlupkothen: Modeling Complex Intertextual Relations



Beispiel: Das *Rheinweinlied* und seine Intertexte

Matthias Claudius (A):

Rheinweinlied

Bekränzt mit Laub den lieben vollen Becher,
Und trinkt ihn fröhlich leer.
In ganz Europia, Ihr Herren Zecher!
Ist solch ein Wein nicht mehr.

Anonym (C):

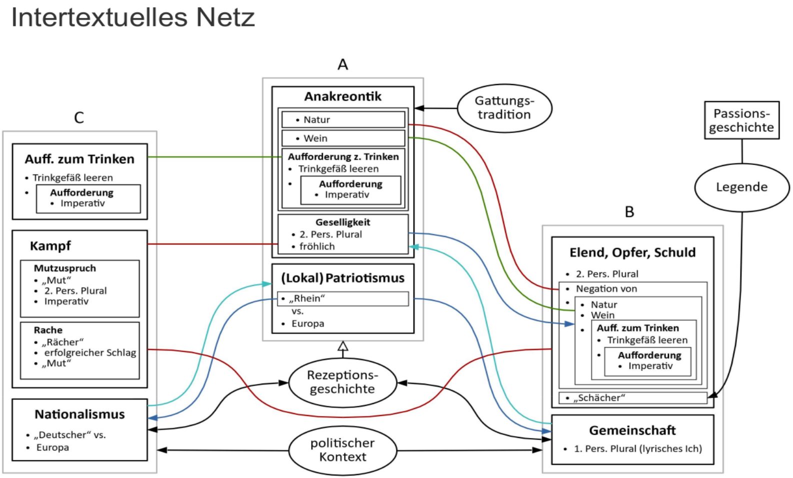
Deutscher Zuruf

Ergreift mit Mut den angefüllten Becher
Und trinkt ihn freudig leer!
Erhoben hat den Arm Europas Rächer,
Und seine Hand traf schwer.

Johann Gottfried Seume (B):

Bekränzt euch nicht, ihr weingelehrten Zecher,
Der Becher bleibe leer;
Denn wir sind nichts als nur Europens Schächer,
Für die ist Wein nicht mehr.

Nantke and Schlupkothen: Modeling Complex Intertextual Relations

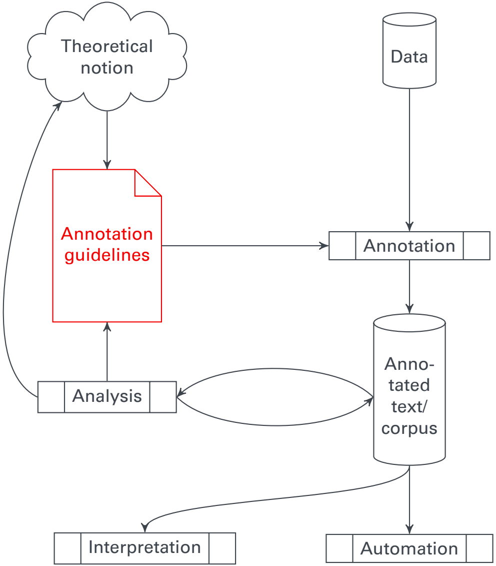

Intertextuelles Netz

Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered

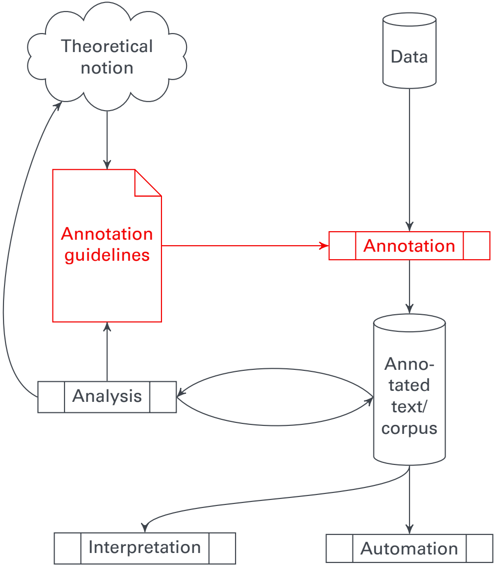
Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered



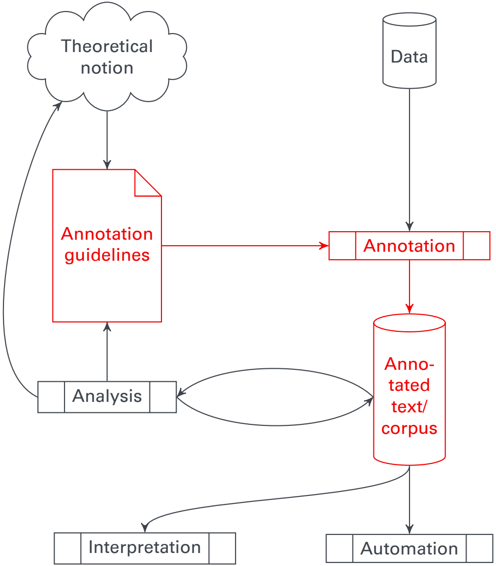
Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered



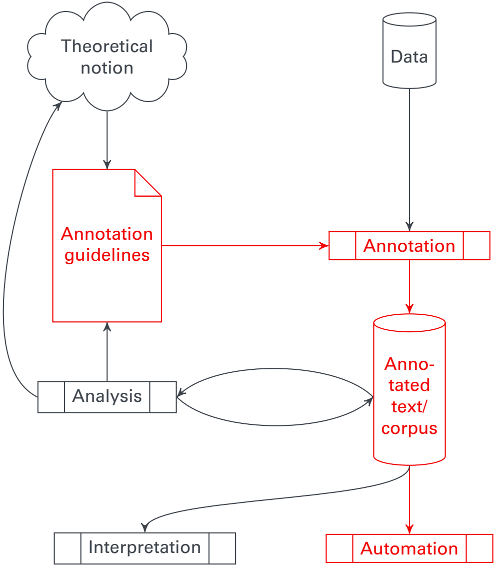
Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered



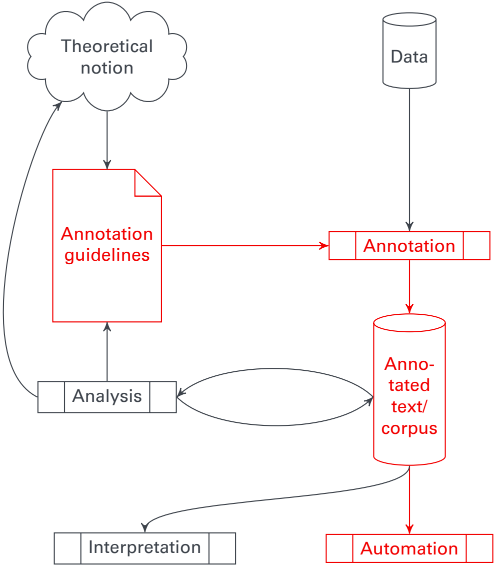
Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered

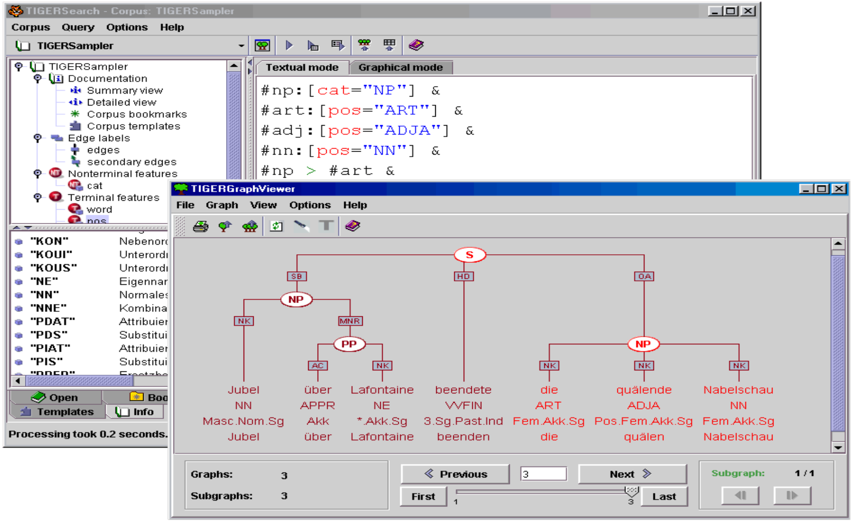


Goals of Annotation

- **Automation-oriented**
  - Create input data for automatic learning procedures
  - Usually uses existing theoretic notions
  - NLP/CL centered
- Examples
  - GRAIN (Schweitzer et al., 2018)
  - Schulz and Kuhn (2016)



POS



```
#np:[cat="NP"]  &
#art:[pos="ART"]  &
#adj:[pos="ADJA"]  &
#nn:[pos="NN"]  &
#np > #art  &
```

"KON"	Nebenord
"KOUI"	Unterordı
"KOUS"	Unterordı
"NE"	Eigennar
"NN"	Normales
"NNE"	Kombina
"PDAT"	Attribuier
"PDS"	Substitui
"PIAT"	Attribuier
"PIS"	Substitui

Processing took 0.2 seconds.

Graphs: 3

Subgraphs: 3

Subgraph: 1 / 1

Discussion

- Annotation in DH
  - Helps disciplines with high theoretic focus
  - Supports intersubjectivity

Discussion

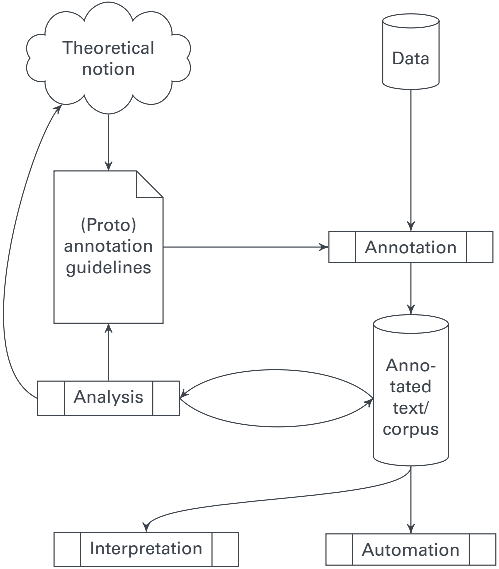
- Annotation in DH
  - Helps disciplines with high theoretic focus
  - Supports intersubjectivity
- Different goals enforce different tools
  - Not entirely clear how tools’ functionality and annotation goals interrelate

Discussion

- Annotation in DH
  - Helps disciplines with high theoretic focus
  - Supports intersubjectivity
- Different goals enforce different tools
  - Not entirely clear how tools' functionality and annotation goals interrelate
- How to treat true disagreements? (Gius and Jacke, 2017)
  - CL usually requires unambiguous annotation data
  - Humanities deal with ambiguity of concepts (different interpretations)
  - How to measure these wanted disagreements?
  - Representation problems
  - Machine learning problems

Conclusion

- Workflow
- Goals
  - Exploratory
  - Conceptualizing
  - Explicating
  - Automation-oriented
- Example projects and areas
  - Pliny (Tool)
  - 3DH (Tool)
  - heureCLÉA (Narratology)
  - QuaDramA (Coreference)
  - Nantke and Schlupkothen (Intertextuality)
  - GRAIN (Part-of-Speech tagging)



Universität Stuttgart

Janis Pagel, Nils Reiter, Ina Rösiger, Sarah Schulz
Institute for Natural Language Processing (IMS), University of Stuttgart

eMail {pageljs,reiterns,roesigia,schulzsh}@ims.uni-stuttgart.de
Phone +49-711-685 813 {89,54,30,94}
Website http://www.ims.uni-stuttgart.de/

References I

Bögel, Thomas et al. (2015). “Collaborative Text Annotation Meets Machine Learning: heureCLÉA, a Digital Heuristic of Narrative”. In: *DHCommons* 1.

Bradley, John (2008). “Thinking about interpretation: Pliny and scholarship in the humanities”. In: *Literary and Linguistic Computing* 23.3, pp. 263–279. DOI: 10.1093/llc/fqn021. URL: http://dx.doi.org/10.1093/llc/fqn021.

Gius, Evelyn and Janina Jacke (2017). “The Hermeneutic Profit of Annotation: On Preventing and Fostering Disagreement in Literary Analysis”. In: *International Journal of Humanities and Arts Computing* 11.2, pp. 233–254. DOI: 10.3366/ijhac.2017.0194. URL: https://doi.org/10.3366/ijhac.2017.0194.

Hovy, Eduard and Julia Lavid (2010). “Towards a ‘Science’ of Corpus Annotation: A New Methodological Challenge for Corpus Linguistics”. In: *International Journal of Translation Studies* 22.1, pp. 13–36.

Kleymann, Rabea, Jan Christoph Meister, and Jan-Erik Stange (Feb. 2018). “Perspektiven kritischer Interfaces für die Digital Humanities im 3DH-Projekt”. In: *Book of Abstracts of DHd 2018*. Cologne, Germany.

Nantke, Julia and Frederik Schlupkothen (2018). “Zwischen Polysemie und Formalisierung: Mehrstufige Modellierung komplexer intertextueller Relationen als Annäherung an ein ‚literarisches’ Semantic Web”. In: *Proceedings of DHd*.

Pustejovsky, James and Amber Stubbs (2012). *Natural Language Annotation for Machine Learning: A Guide to Corpus-Building for Applications*. Sebastopol, Boston, Farnham: O’Reilly Media. ISBN: 9781449306663.

Rösiger, Ina, Sarah Schulz, and Nils Reiter (2018). “Towards Coreference for Literary Text: Analyzing Domain-Specific Phenomena”. In: *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Santa Fe, USA.

Schulz, Sarah and Jonas Kuhn (2016). “Learning from Within? Comparing PoS Tagging Approaches for Historical Text.”. In: *LREC*. Ed. by Nicoletta Calzolari et al. European Language Resources Association (ELRA). URL: http://dblp.uni-trier.de/db/conf/lrec/lrec2016.html#SchulzK16.

Schweitzer, Katrin et al. (2018). “German Radio Interviews: The GRAIN Release of the SFB732 Silver Standard Collection”. In: *Proceedings of the 11th International Conference on Language Resources and Evaluation*. LREC 2018.